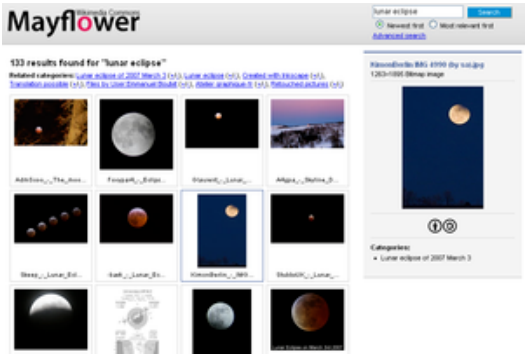
Web search engine

A **web search engine** or **Internet search engine** is a software system that is designed to carry out **web search** (**Internet search**), which means to search the World Wide Web in a systematic way for particular information specified in a web search query. The search results are generally presented in a line of results, often referred to as search engine results pages (SERPs). The information may be a mix of web pages, images, videos, infographics, articles, research papers and other types of files. Some search engines also mine data available in databases or open directories. Unlike web directories, which are maintained only by human editors, search engines also maintain real-time information by running an algorithm on a web crawler. Internet content that is not capable of being searched by a web search engine is generally described as the deep web.



The results of a search for the term "lunar eclipse" in a web-based image search engine

Contents

History

Approach

Market share

- Market share in February 2019
- East Asia and Russia
- Europe

Search engine bias

Customized results and filter bubbles

Christian, Islamic and Jewish search engines

Search engine submission

See also

References

Further reading

External links

- User's navigation

History

Internet search engines themselves predate the debut of the Web in December 1990. The Who is user search dates back to 1982[1] and the Knowbot Information Service multi-network user search was first implemented in 1989.[2] The first well documented search engine that searched content files, namely FTP files was Archie, which debuted on 10 September 1990.[3]

Prior to September 1993, the World Wide Web was entirely indexed by hand. There was a list of webservers edited by Tim Berners-Lee and hosted on the CERN webserver. One snapshot of the list in 1992 remains,[4] but as more and more web servers went online the central list could no longer keep up. On the NCSA site, new servers were announced under the title "What's New!"[5]

The first tool used for searching content (as opposed to users) on the Internet was Archie.[6] The name stands for "archive" without the "v". It was created by Alan Emtage, Bill Heelan and J. Peter Deutsch, computer science students at McGill University in Montreal, Quebec, Canada. The program downloaded the directory listings of all the files located on public anonymous FTP (File Transfer Protocol) sites, creating a searchable database of file names; however, Archie Search Engine did not index the contents of these sites since the amount of data was so limited it could be readily searched manually.

The rise of Gopher (created in 1991 by Mark McCahill at the University of Minnesota) led to two new search programs, Veronica and Jughead. Like Archie, they searched the file names and titles stored in Gopher index systems. Veronica (*V*ery *E*asy *R*odent-*O*riented *N*et-wide *I*ndex to *C*omputerized *A*rchives) provided a keyword search of most Gopher menu titles in the entire Gopher listings. Jughead (*J*onzy's *U*niversal *G*opher *H*ierarchy *E*xcavation *A*nd *D*isplay) was a tool for obtaining menu information from specific Gopher servers. While the name of the search engine "Archie Search Engine" was not a reference to the Archie comic book series, "Veronica" and "Jughead" are characters in the series, thus referencing their predecessor

In the summer of 1993, no search engine existed for the web, though numerous specialized catalogues were maintained by hand. Oscar Nierstrasz at the University of Geneva wrote a series of Perl scripts that periodically mirrored these pages and rewrote them into a standard format. This formed the basis for W3Catalog, the web's first primitive search engine, released on September 2, 1993.[7]

In June 1993, Matthew Gray, then at MIT, produced what was probably the first web robot, the Perl-based World Wide Web Wanderer, and used it to generate an index called 'Wandex'. The purpose of the Wanderer was to measure the size of the World Wide Web, which it did until late 1995. The web's second search engine Aliweb appeared in November 1993. Aliweb did not use a web robot, but instead depended on being notified by website administrators of the existence at each site of an index file in a particular format.

JumpStation (created in December 1993[8] by Jonathon Fletcher) used a web robot to find web pages and to build its index, and used a web form as the interface to its query program. It was thus the first WWW resource-discovery tool to combine the three essential features of a web search engine (crawling, indexing, and searching) as described below. Because of the limited resources available on the platform it ran on, its indexing and hence searching were limited to the titles and headings found in the web pages the crawler encountered.

One of the first "all text" crawler-based search engines was WebCrawler, which came out in 1994. Unlike its predecessors, it allowed users to search for any word in any webpage, which has become the standard for all major search engines since. It was also the search engine that was widely known by the public. Also in 1994, Lycos (which started at Carnegie Mellon University) was launched and became a major commercial endeavor.

Soon after, many search engines appeared and vied for popularity. These included Magellan, Excite, Infoseek, Inktomi, Northern Light, and AltaVista. Yahoo! was among the most popular ways for people to find web pages of interest, but its search function operated on its web directory, rather than its full-text copies of web pages. Information seekers could also browse the directory instead of doing a keyword-based search.

In 1996, Netscape was looking to give a single search engine an exclusive deal as the featured search engine on Netscape's web browser. There was so much interest that instead Netscape struck deals with five of the major search engines: for $5 million a year, each search engine would be in rotation on the Netscape search engine page. The five engines were Yahoo!, Magellan, Lycos, Infoseek, and Excite.[9][10]

Google adopted the idea of selling search terms in 1998, from a small search engine company named goto.com. This move had a significant effect on the SE business, which went from struggling to one of the most profitable businesses in the Internet.[11]

Search engines were also known as some of the brightest stars in the Internet investing frenzy that occurred in the late 1990s.[12] Several companies entered the market spectacularly, receiving record gains during their initial public offerings. Some have taken down their public search engine, and are marketing enterprise-only editions, such as Northern Light. Many search engine companies were caught up in the dot-com bubble, a speculation-driven market boom that peaked in 1999 and ended in 2001.

- Around 2000, Google's search engine rose to prominence.[13] The company achieved better results for many searches with an innovation called PageRank, as was explained in the paper *Anatomy of a Search Engine* written by Sergey Brin and Larry Page, the later founders of Google.[14] This iterative algorithm ranks web pages based on the number and PageRank of other web sites and pages that link there, on the premise that good or desirable pages are linked to more than others. Google also maintained a minimalist interface to its search engine. In contrast, many of its competitors embedded a search engine in a web portal. In fact, Google search engine became so popular that spoof engines emerged such as Mystery Seeker.

By 2000, Yahoo! was providing search services based on Inktomi's search engine. Yahoo! acquired Inktomi in 2002, and Overture (which owned AlltheWeb and AltaVista) in 2003. Yahoo! switched to Google's search engine until 2004, when it launched its own search engine based on the combined technologies of its acquisitions.

Timeline (full list)

Year	Engine	Current status
1993	W3Catalog	Inactive
	Aliweb	Inactive
	JumpStation	Inactive
	WWW Worm	Inactive
1994	WebCrawler	Active
	Go.com	Inactive, redirects to Disney
	Lycos	Active
	Infoseek	Inactive, redirects to Disney
1995	Daum	Active
	Magellan	Inactive
	Excite	Active
	SAPO	Active
	MetaCrawler	Active
	Yahoo!	Active, Launched as a directory
	AltaVista	Inactive, acquired by Yahoo! in 2003, since 2013 redirects to Yahoo!
1996	Dogpile	Active, Aggregator
	Inktomi	Inactive, acquired by Yahoo!
	HotBot	Active
	Ask Jeeves	Active (rebranded ask.com)
1997	AOL NetFind	Active (rebranded AOL Search since 1999)
	Northern Light	Inactive
	Yandex	Active
1998	Google	Active
	Ixquick	Active as Startpage.com
	MSN Search	Active as Bing
	empas	Inactive (merged with

Microsoft first launched MSN Search in the fall of 1998 using search results from Inktomi. In early 1999 the site began to display listings from Looksmart, blended with results from Inktomi. For a short time in 1999, MSN Search used results from AltaVista instead. In 2004, Microsoft began a transition to its own search technology powered by its own web crawler (called msnbot).

Microsoft's rebranded search engine, Bing, was launched on June 1, 2009. On July 29, 2009, Yahoo! and Microsoft finalized a deal in which Yahoo! Search would be powered by Microsoft Bing technology.

As of 2018, active search engine crawlers include that of Google, Bing, Gigablast, Mojeek, Baidu and Yandex.

Approach

A search engine maintains the following processes in near real time:

1. Web crawling
2. Indexing
3. Searching[15]

Web search engines get their information by web crawling from site to site. The "spider" checks for the standard filename *robots.txt*, addressed to it, before sending certain information back to be indexed depending on many factors, such as the titles, page content, JavaScript, Cascading Style Sheets (CSS), headings, as evidenced by the standard HTML markup of the informational content, or its metadata in HTML meta tags. "[N]o web crawler may actually crawl the entire reachable web. Due to infinite websites, spider traps, spam, and other exigencies of the real web, crawlers instead apply a crawl policy to determine when the crawling of a site should be deemed sufficient. Some sites are crawled exhaustively while others are crawled only partially".[16]

Indexing means associating words and other definable tokens found on web pages to their domain names and HTML-based fields. The associations are made in a public database, made available for web search queries. A query from a user can be a single word. The index helps find information relating to the query as quickly as possible.[15] Some of the techniques for indexing, and caching are trade secrets, whereas web crawling is a straightforward process of visiting all sites on a systematic basis.

Between visits by the *spider*, the cached version of page (some or all the content needed to render it) stored in the search engine working memory is quickly sent to an inquirer. If a visit is overdue, the search engine can just act as a web proxy instead. In this case the page may differ from the search terms indexed.[15] The cached page holds the appearance of the version whose words were indexed, so a cached version of a page can be useful to the web site when the actual page has been lost, but this problem is also considered a mild form of linkrot.

Typically when a user enters a query into a search engine it is a few keywords.[17] The index already has the names of the sites containing the keywords, and these are instantly obtained from the index. The real processing load is in generating the web pages that are the search results list: Every page in the entire list must be weighted according to information in the indexes.[15] Then the top search result item requires the lookup, reconstruction, and markup of the *snippets* showing the context of the keywords matched. These are only part of the processing each search results web page requires, and further pages (next to the top) require more of this post processing.

Beyond simple keyword lookups, search engines offer their own GUI- or command-driven operators and search parameters to refine the search results. These provide the necessary controls for the user engaged in the feedback loop users create by *filtering* and *weighting* while refining the search results, given the initial pages of the first search results. For example, from 2007 the Google.com search engine has allowed one to *filter* by date by clicking "Show search tools" in the leftmost column of the initial search results page, and then selecting the desired date range.[18] It's also possible to *weight* by date because each page has a modification time. Most search engines support the use of the boolean operators AND, OR and NOT to help end users refine the search query. Boolean operators are for literal searches that allow the user to refine and extend the terms of the

		NATE)
1999	AlltheWeb	Inactive (URL redirected to Yahoo!)
	GenieKnows	Active, rebranded Yellowee.com
	Naver	Active
	Teoma	Inactive,
2000	Baidu	Active
	Exalead	Active
	Gigablast	Active
2001	Kartoo	Inactive
2003	Info.com	Active
	Scroogle	Inactive
2004	Yahoo! Search	Active, Launched own web search (see Yahoo! Directory, 1995)
	A9.com	Inactive
	Clusty	Active (as Yippy)
	Mojeek	Active
	Sogou	Active
2005	SearchMe	Inactive
	KidzSearch	Active, Google Search
2006	Soso	Inactive, merged with Sogou
	Quaero	Inactive
	Search.com	Active
	ChaCha	Inactive
	Ask.com	Active
	Live Search	Active as Bing, rebranded MSN Search
2007	wikiseek	Inactive
	Sproose	Inactive
	Wikia Search	Inactive
	Blackle.com	Active, Google Search
2008	Powerset	Inactive (redirects to Bing)
	Picollator	Inactive
	Viewzi	Inactive

search. The engine looks for the words or phrases exactly as entered. Some search engines provide an advanced feature called proximity search, which allows users to define the distance between keywords.[15] There is also concept-based searching where the research involves using statistical analysis on pages containing the words or phrases you search for. As well, natural language queries allow the user to type a question in the same form one would ask it to a human.[19] A site like this would be ask.com![20]

The usefulness of a search engine depends on the relevance of the *result set* it gives back. While there may be millions of web pages that include a particular word or phrase, some pages may be more relevant, popular, or authoritative than others. Most search engines employ methods to rank the results to provide the "best" results first. How a search engine decides which pages are the best matches, and what order the results should be shown in, varies widely from one engine to another.[15] The methods also change over time as Internet usage changes and new techniques evolve. There are two main types of search engine that have evolved: one is a system of predefined and hierarchically ordered keywords that humans have programmed extensively. The other is a system that generates an "inverted index" by analyzing texts it locates. This first form relies much more heavily on the computer itself to do the bulk of the work.

Most Web search engines are commercial ventures supported by advertising revenue and thus some of them allow advertisers to have their listings ranked higher in search results for a fee. Search engines that do not accept money for their search results make money by running search related ads alongside the regular search engine results. The search engines make money every time someone clicks on one of these ads.[21]

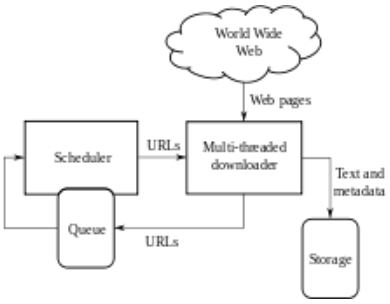
Market share

Google is the world's most popular search engine, with a market share of 92.86 percent as of February, 2019.[22]

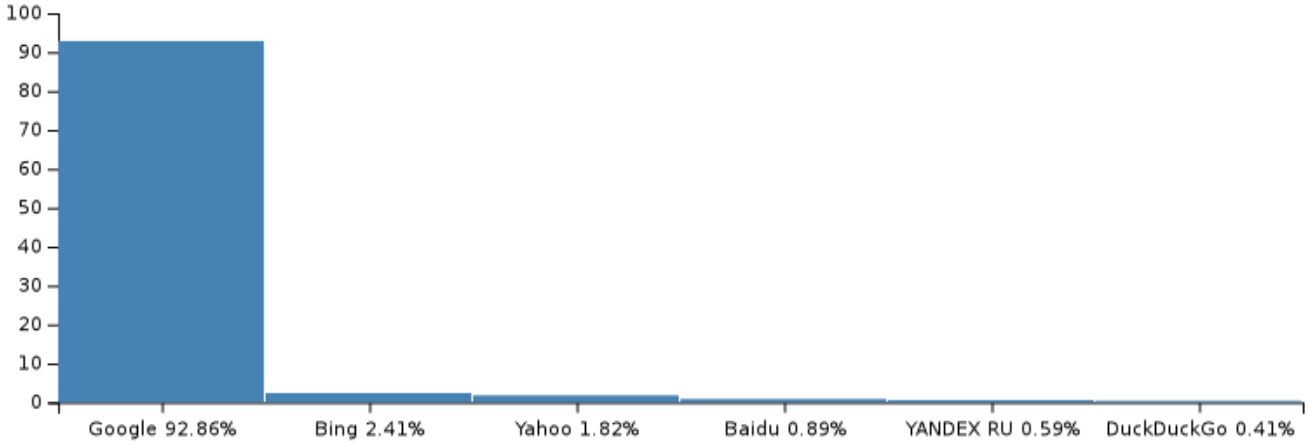
The world's most popular search engines are:

Market share in February 2019

	Boogami	Inactive
	LeapFish	Inactive
	Forestle	Inactive (redirects to Ecosia)
	DuckDuckGo	Active
2009	Bing	Active, rebranded Live Search
	Yebol	Inactive
	Mugurdy	Inactive due to a lack of funding
	Scout (Goby)	Active
	NATE	Active
	Ecosia	Active
	Startpage.com	Active, sister engine of Ixquick
2010	Blekko	Inactive, sold to IBM
	Cuil	Inactive
	Yandex (English)	Active
	Parsijoo	Active
2011	YaCy	Active, P2P web search engine
2012	Volunia	Inactive
2013	Qwant	Active
2014	Egerin	Active, Kurdish / Sorani search engine
	Swisscows	Active
2015	Yooz	Active
	Cliqz	Active, Browser integrated search engine
2016	Search Encrypt	Active
	Kiddle	Active, Google Search



High-level architecture of a standard Web crawler



East Asia and Russia

In some East Asian countries and Russia, Google is not the most popular search engine.

In Russia, Yandex commands a marketshare of 61.9 percent, compared to Google's 28.3 percent.[23] In China, Baidu is the most popular search engine.[24] South Korea's homegrown search portal, Naver, is used for 70 percent of online searches in the country.[25] Yahoo! Japan and Yahoo! Taiwan are the most popular avenues for Internet searches in Japan and Taiwan, respectively.[26]

Europe

Most countries' markets in Western Europe are dominated by Google, except for Czech Republic, where Seznam is a strong competitor.[27]

Search engine bias

Although search engines are programmed to rank websites based on some combination of their popularity and relevancy, empirical studies indicate various political, economic, and social biases in the information they provide.[28][29] and the underlying assumptions about the technology.[30] These biases can be a direct result of economic and commercial processes (e.g., companies that advertise with a search engine can become also more popular in its organic search results), and political processes (e.g., the removal of search results to comply with local laws).[31] For example, Google will not surface certain neo-Nazi websites in France and Germany where Holocaust denial is illegal.

Biases can also be a result of social processes, as search engine algorithms are frequently designed to exclude non-normative viewpoints in favor of more "popular" results.[32] Indexing algorithms of major search engines skew towards coverage of U.S.-based sites, rather than websites from non-U.S. countries.[29]

Google Bombing is one example of an attempt to manipulate search results for political, social or commercial reasons.

Several scholars have studied the cultural changes triggered by search engines,[33] and the representation of certain controversial topics in their results, such as terrorism in Ireland,[34] climate change denial,[35] and conspiracy theories.[36]

Customized results and filter bubbles

Many search engines such as Google and Bing provide customized results based on the user's activity history. This leads to an effect that has been called a filter bubble. The term describes a phenomenon in which websites use algorithms to selectively guess what information a user would like to see, based on information about the user (such as location, past click behaviour and search history). As a result, websites tend to show only information that agrees with the user's past viewpoint. This puts the user in a state of intellectual isolation without contrary information. Prime examples are Google's personalized search results and Facebook's personalized news stream. According to Eli Pariser, who coined the term, users get less exposure to conflicting viewpoints and are isolated intellectually in their own informational bubble. Pariser related an example in which one user searched Google for "BP" and got investment news about British Petroleum while another searcher got information about the Deepwater Horizon oil spill and that the two search results pages were "strikingly different".[37][38][39] The bubble effect may have negative implications for civic discourse, according to Pariser.[40] Since this problem has been identified, competing search engines have emerged that seek to avoid this problem by not tracking or "bubbling" users, such as DuckDuckGo. Other scholars do not share Pariser's view, finding the evidence in support of his thesis unconvincing.[41]

Christian, Islamic and Jewish search engines

The global growth of the Internet and electronic media in the Arab and Muslim World during the last decade has encouraged Islamic adherents in the Middle East and Asian sub-continent, to attempt their own search engines, their own filtered search portals that would enable users to perform safe searches. More than usual *safe search* filters, these Islamic web portals categorizing websites into being either "halal" or "haram", based on modern, expert, interpretation of the "Law of Islam". ImHalal came online in September 2011. Halalgoogling came online in July 2013. These use haram filters on the collections from Google and Bing (and others).[42]

While lack of investment and slow pace in technologies in the Muslim World has hindered progress and thwarted success of an Islamic search engine, targeting as the main consumers Islamic adherents, projects like Muxlim, a Muslim lifestyle site, did receive millions of dollars from investors like Rite Internet Ventures, and it also faltered. Other religion-oriented search engines are Jewogle, the Jewish version of Google,[43] and SeekFind.org, which is Christian. SeekFind filters sites that attack or degrade their faith.[44]

Search engine submission

Web search engine submission is a process in which a webmaster submits a website directly to a search engine. While search engine submission is sometimes presented as a way to promote a website, it generally is not necessary because the major search engines use web crawlers, that will eventually find most web sites on the Internet without assistance. They can either submit one web page at a time, or they can submit the entire site using a sitemap, but it is normally only necessary to submit the home page of a web site as search engines are able to crawl a well designed website. There are two remaining reasons to submit a web site or web page to a search engine: to add an entirely new web site without waiting for a search engine to discover it, and to have a web site's record updated after a substantial redesign.

Some search engine submission software not only submits websites to multiple search engines, but also add links to websites from their own pages. This could appear helpful in increasing a website's ranking, because external links are one of the most important factors determining a website's ranking. However, John Mueller of Google has stated that this "can lead to a tremendous number of unnatural links for your site" with a negative impact on site ranking.[45]

See also

-
- Comparison of web search engines
- Information retrieval
- List of search engines
- Question answering
- Google effect
- Use of web search engines in libraries
- Semantic Web
- Spell checker
- Web development tools
- Search engine manipulation effect
- Search engine privacy

References

1. "RFC 812 - NICNAME/WHOIS" (https://tools.ietf.org/html/rfc812). *ietf.org*.
2. "Knowbot programming: System support for mobile agents" (http://www.cnri.reston.va.us/home/koe/iwoos-full.html). *cnri.reston.va.us*.
3. Deutsch, Peter (September 11, 1990). "[next] An Internet archive server server (was about Lisp)" (https://groups.google.com/forum/#!msg/comp.archives/LWVA50W8BKk/wyRbF_IDc6cJ). *groups.google.com*. Retrieved 2017-12-29.

4. "World-Wide Web Servers" (http://www.w3.org/History/19921103-hypertext/hypertext/DataSources/WWW/Servers.html). W3.org. Retrieved 2012-05-14.
5. "What's New! February 1994" (http://home.mcom.com/home/whatsnew/whats_new_0294.html). Home.mcom.com. Retrieved 2012-05-14.
6. "Internet History - Search Engines" (from Search Engine Watch), Universiteit Leiden, Netherlands, September 2001, web: LeidenU-Archie (http://www.internethistory.leidenuniv.nl/index.php3?c=7).
7. Oscar Nierstrasz (2 September 1993). "Searchable Catalog of WWW Resources (experimental)" (http://groups.google.com/group/comp.infosystems.www/browse_thread/thread/2176526a36dc8bd3/2718fd17812937ac?hl=en&lnk=gst&q=Oscar+Nierstrasz#2718fd17812937ac).
8. "Archive of NCSA what's new in December 1993 page" (https://web.archive.org/web/20010620073530/http://archive.ncsa.uiuc.edu/SDG/Software/Mosaic/Docs/old-whats-new/whats-new-1293.html). 2001-06-20. Archived from the original (http://archive.ncsa.uiuc.edu/SDG/Software/Mosaic/Docs/old-whats-new/whats-new-1293.html) on 2001-06-20. Retrieved 2012-05-14.
9. "Yahoo! And Netscape Ink International Distribution Deal" (http://files.shareholder.com/downloads/YHOO/701084386x0x27155/9a3b5ed8-9e84-4cba-a1e5-77a3dc606566/YHOO_News_1997_7_8_General.pdf) (PDF).
10. "Browser Deals Push Netscape Stock Up 7.8%" (http://articles.latimes.com/1996-04-01/business/fi-53780_1_netscape-home). Los Angeles Times. 1 April 1996
11. Pursel, Bart. "Search Engines" (https://psu.pb.unizin.org/ist110/chapter/2-1-search-engines/). *Penn State Pressbooks*. Retrieved February 20, 2018.
12. Gandal, Neil (2001). "The dynamics of competition in the internet search engine market". *International Journal of Industrial Organization*. **19** (7): 1103–1117. doi:10.1016/S0167-7187(01)00065-0 (https://doi.org/10.1016%2FS0167-7187%2801%2900065-0)
13. "Our History in depth" (http://www.google.com/about/company/history/). W3.org. Retrieved 2012-10-31.
14. Brin, Sergey; Page, Larry. "The Anatomy of a Large-Scale Hypertextual Web Search Engine" (http://ilpubs.stanford.edu:8090/361/1/1998-8.pdf) (PDF).
15. Jawadekar, Waman S (2011), "8. Knowledge Management: Tools and Technology" (https://books.google.com/books?id=XmGx4J9daUMC&pg=PA278), *Knowledge Management: Text & Cases*, New Delhi: Tata McGraw-Hill Education Private Ltd, p. 278, ISBN 978-0-07-07-0086-4, retrieved November 23, 2012
16. Dasgupta, Anirban; Ghosh, Arpita; Kumar, Ravi; Olston, Christopher; Pandey, Sandeep; and Tomkins, Andrew. *The Discoverability of the Web*. http://www.arpitaghosh.com/papers/discoverability.pdf
17. Jansen, B. J., Spink, A., and Saracevic, T. 2000. Real life, real users, and real needs: A study and analysis of user queries on the web. Information Processing & Management (https://faculty.ist.psu.edu/jjansen/academic/pubs/jansen_real_life_real_users_and_real_needs.pdf). 36(2), 207-227.
18. Chitu, Alex (August 30, 2007). "Easy Way to Find Recent Web Pages" (http://googlesystem.blogspot.com/2007/08/easy-way-to-find-recent-web-pages.html). *Google Operating System*. Retrieved 22 February 2015.
19. "Versatile question answering systems: seeing in synthesis (https://www.academia.edu/2475776/Versatile_question_answering_systems_seeing_in_synthesis)", Mittal et al., IJIIDS, 5(2), 119-142, 2011.
20. http://www.ask.com. Retrieved 10 September 2015.
21. "how search engine works?" (http://globalforumonline.com/detail/how-does-search-engine-works/). GFO = 26 June 2018.
22. "Search Engine Market Share Worldwide" (http://gs.statcounter.com/search-engine-market-share). *statcounter*. gs.statcounter.com. Retrieved 24 February 2019.
23. "Live Internet - Site Statistics" (http://www.liveinternet.ru/stat/ru/searches.html?slice=ru;period=week). Live Internet. Retrieved 2014-06-04.
24. Arthur, Charles (2014-06-03). "The Chinese technology companies poised to dominate the world" (https://www.theguardian.com/world/2014/jun/03/chinese-technology-companies-huawei-dominate-world). The Guardian. Retrieved 2014-06-04.
25. "How Naver Hurts Companies' Productivity" (https://blogs.wsj.com/korearealtime/2014/05/21/how-naver-hurts-companies-productivity/). The Wall Street Journal. 2014-05-21 Retrieved 2014-06-04.
26. "Age of Internet Empires" (http://geography.oii.ox.ac.uk/?page=age-of-internet-empires). Oxford Internet Institute. Retrieved 2014-06-0 Check date values in: `|accessdate=` (help)
27. Seznam Takes on Google in the Czech Republic (http://www.doz.com/search-engine/seznam-search-engine). Doz.
28. Segev, El (2010). Google and the Digital Divide: The Biases of Online Knowledge, Oxford: Chandos Publishing.
29. Vaughan, Liwen; Mike Thelwall (2004). "Search engine coverage bias: evidence and possible causes". *Information Processing & Management*. **40** (4): 693–707. CiteSeerX 10.1.1.65.5130 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.65.5130) doi:10.1016/S0306-4573(03)00063-3 (https://doi.org/10.1016%2FS0306-4573%2803%2900063-3)
30. Jansen, B. J. and Rieh, S. (2010) The Seventeen Theoretical Constructs of Information Searching and Information Retrieval (https://faculty.ist.psu.edu/jjansen/academic/jansen_theoretical_constructs.pdf) Journal of the American Society for Information Sciences and Technology. 61(8), 1517-1534.
31. Berkman Center for Internet & Society (2002), "Replacement of Google with Alternative Search Systems in China: Documentation and Screen Shots" (http://cyber.law.harvard.edu/filtering/china/google-replacements/). Harvard Law School.
32. Introna, Lucas; Helen Nissenbaum (2000). "Shaping the Web: Why the Politics of Search Engines Matters". *The Information Society: An International Journal* **16** (3): 169–185. CiteSeerX 10.1.1.24.8051 (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.24.8051). doi:10.1080/01972240050133634 (https://doi.org/10.1080%2F01972240050133634)

33. Hillis, Ken; Petit, Michael; Jarrett, Kylie (2012-10-12). *Google and the Culture of Search* (https://books.google.com/books?id=R7Lzp7apkJgC). Routledge. ISBN 9781136933066.
34. Reilly, P. (2008-01-01). Spink, Prof Dr Amanda; Zimmer, Michael, eds. *'Googling' Terrorists: Are Northern Irish Terrorists Visible on Internet Search Engines?*. Information Science and Knowledge Management. **14**. Springer Berlin Heidelberg. pp. 151–175. doi:10.1007/978-3-540-75829-7_10 (https://doi.org/10.1007%2F978-3-540-75829-7_10). ISBN 978-3-540-75828-0.
35. Hiroko Tabuchi, "How Climate Change Deniers Rise to the Top in Google Searches (https://www.nytimes.com/2017/12/29/climate/google-search-climate-change.html)", The New York Times, Dec. 29, 2017. Retrieved November 14, 2018.
36. Ballatore, A (2015). "Google chemtrails: A methodology to analyze topic representation in search engines" (http://firstmonday.org/ojs/index.php/fm/article/view/5597) *First Monday*. **20** (7). doi:10.5210/fm.v20i7.5597 (https://doi.org/10.5210%2Ffm.v20i7.5597).
37. Parramore, Lynn (10 October 2010). "The Filter Bubble" (https://www.theatlantic.com/daily-dish/archive/2010/10/the-filter-bubble/181427/). *The Atlantic*. Retrieved 2011-04-20. "Since Dec. 4, 2009, Google has been personalized for everyone. So when I had two friends this spring Google "BP," one of them got a set of links that was about investment opportunities in BP. The other one got information about the oil spill..."
38. Weisberg, Jacob (10 June 2011). "Bubble Trouble: Is Web personalization turning us into solipsistic twits?" (http://www.slate.com/id/2296633/). *Slate*. Retrieved 2011-08-15.
39. Gross, Doug (May 19, 2011). "What the Internet is hiding from you" (http://edition.cnn.com/2011/TECH/web/05/19/online.privacy.pariser/). *CNN*. Retrieved 2011-08-15. "I had friends Google BP when the oil spill was happening. These are two women who were quite similar in a lot of ways. One got a lot of results about the environmental consequences of what was happening and the spill. The other one just got investment information and nothing about the spill at all."
40. Zhang, Yuan Cao; Séaghdha, Diarmuid Ó; Quercia, Daniele; Jambor, Tamas (February 2012). "Auralist: Introducing Serendipity into Music Recommendation" (http://www-typo3.cs.ucl.ac.uk/fileadmin/UCL-CS/research/Research_Notes/RN_11_21.pdf) (PDF). *ACM Wsdm*.
41. O'Hara, K. (2014-07-01). "In Worship of an Echo". *IEEE Internet Computing* **18** (4): 79–83. doi:10.1109/MIC.2014.71 (https://doi.org/10.1109%2FMIC.2014.71). ISSN 1089-7801 (https://www.worldcat.org/issn/1089-7801).
42. "New Islam-approved search engine for Muslims" (http://news.msn.com/science-technology/new-islam-approved-search-engine-for-muslims). News.msn.com. Retrieved 2013-07-11.
43. "Jewogle - FAQ" (http://www.jewogle.com/faq/).
44. "Halalgoogling: Muslims Get Their Own "sin free" Google; Should Christians Have Christian Google? - Christian Blog" (http://allchristiannews.com/halalgoogling-muslims-get-their-own-sin-free-google-should-christians-have-christian-google/). *Christian Blog*. 2013-07-25.
45. Schwartz, Barry (2012-10-29). "Google: Search Engine Submission Services Can Be Harmful" (https://www.seroundtable.com/search-engine-submission-google-15906.html). *Search Engine Roundtable*. Retrieved 2016-04-04.

Further reading

- Steve Lawrence; C. Lee Giles (1999). "Accessibility of information on the web". *Nature*. **400** (6740): 107–9. Bibcode:1999Natur.400..107L. doi:10.1038/21987. PMID 10428673.
- Bing Liu (2007), *Web Data Mining: Exploring Hyperlinks, Contents and Usage Data*. Springer, ISBN 3-540-37881-2
- Bar-Ilan, J. (2004). The use of Web search engines in information science research. ARIST, 38, 231-288.
- Levene, Mark (2005). *An Introduction to Search Engines and Web Navigation*. Pearson.
- Hock, Randolph (2007). *The Extreme Searcher's Handbook*. ISBN 978-0-910965-76-7
- Javed Mostafa (February 2005). "Seeking Better Web Searches". *Scientific American*. **292** (2): 66–73. doi:10.1038/scientificamerican0205-66
- Ross, Nancy; Wolfram, Dietmar (2000). "Enduser searching on the Internet: An analysis of term pair topics submitted to the Excite search engine". *Journal of the American Society for Information Science*. **51** (10): 949–958. doi:10.1002/1097-4571(2000)51:10<949::AID-ASI70>3.0.CO;2-5
- Xie, M.; et al. (1998). "Quality dimensions of Internet search engines". *Journal of Information Science* **24** (5): 365–372. doi:10.1177/016555159802400509
- *Information Retrieval: Implementing and Evaluating Search Engines*. MIT Press. 2010.

External links

- Search Engines at Curlie
- "Book Search Engine".
- "List of 42 Advanced Search Operators." Archived from the original on Sep 12, 2018.

User's navigation

- "Delete Recent Searches, on browser and from Web Sites". *WikiHow* (in English, Spanish, German, French, and Italian).
- "Google Knowledge Graph Entities Search Engine" (in English and Italian).

Retrieved from 'https://en.wikipedia.org/w/index.php?title=Web_search_engine&oldid=891972418'

This page was last edited on 11 April 2019, at 10:27 (UTC).

Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc, a non-profit organization.